DOE/NSF U.S. LHC Program Office

Report on the Joint DOE/NSF Review of U.S. LHC Computing

Fermilab
13-16 January 2004

Report to the
US Department of Energy
and the
National Science Foundation

Executive Summary

ATLAS and CMS will be large, general-purpose detectors for observation of very high energy proton-proton collisions at the Large Hadron Collider (LHC). This facility is currently under construction at CERN, the European Laboratory for Particle Physics, near Geneva, Switzerland. In order to reap the scientific benefits of over $0.5 billion of U.S. investment in constructing the LHC and its detectors, the LHC software and computing projects must be successful in empowering the U.S. community of scientists and students to perform frontline physics analysis. The U.S. Computing Projects must not only contribute an appropriate share of computing resources and software effort, but they must also ensure that U.S. physicists will be able to fully and immediately perform analysis of the data and contribute meaningfully to the research work of the physics groups.

A peer review of the U.S. LHC Software and Computing Project was held on 13-16 January 2004 at the Fermi National Accelerator Laboratory in Batavia, Illinois. The primary purpose of the review was to assess the current understanding of the scope, cost and schedule for the US LHC Software and Computing (S&C) projects and the operation of their management structures. Both US ATLAS and US CMS were asked to provide self-consistent project plans targeted to the funding guidance provided from the U.S. Department of Energy (DOE) and National Science Foundation (NSF) and to present contingency plans that could deal with funding levels as much as 20% below the guidance. Because of the dynamic nature of software and computing, the review emphasized plans for the next two years, and concentrated on issues arising from major changes during the past year that have caused changes to the initial baselines adopted at the November 2001 review. These include changes to the LHC schedule, the advent of the LHC Computing Grid (LCG) project, and new funding guidelines from the agencies. In addition, the review addressed how external efforts, particularly from grid R&D and deployment projects, are integrated into the LHC Software and Computing projects. Lastly, since the projects have been operating for roughly two years since the initial baselining, an assessment of technical progress was made. The complete charge to the review panel is shown in Appendix A. Appendices B and C contain the list of members of the review committee and the review agenda, respectively.

The expert reviewers provided comments throughout the review directly to the U.S. LHC collaboration members as well as to the DOE and NSF representatives present. They asked questions of the collaborations both orally and in writing, and deliberated issues and impressions in executive session with only agency representatives present. These comments and those provided in writing by the reviewers, containing their observations, assessments and recommendations, form the basis of this report.

Several overriding themes were relevant to the entire review of both of the experiments. First, the review was somewhat overwhelmed by the announcement of the likelihood that the funding agencies might be unable to fulfill the previously announced funding guidelines, and both experiments were asked on rather short notice to describe how they might respond to such funding shortfalls. While both experiments were able to describe in general terms what the effects of such shortfalls would be, it was clear to the committee that this was a very different situation than previous cases of delayed or reduced funding. The Software and Computing projects have a history of coping with reduced funding by delaying hires and purchases and by somewhat reducing scope, but the effect of these decisions have been ameliorated by concomitant delays in the LHC turn-on schedule. Since it appears that the LHC

schedule is unlikely to have any further major slippages, the current round of potential funding shortfalls will have serious impacts on the U.S. role in the experiments.

In particular, the committee found that the proposed budget cuts would:

- curtail U.S. ability to do its share in upcoming data challenges: the cuts would further defer hardware acquisitions at both Tier 1 and Tier 2 U.S. centers, further eroding the amount of hardware available in the U.S. to be applied to the international data challenges. More importantly, the cuts would severely limit manpower at the Tier 1 centers, further hurting our ability to fully participate in the challenges;
- compromise the ability of U.S. university physicists to do data analysis: budget cuts would limit the ability of U.S. software engineers to support development of physics software, especially at U.S. universities, and the lack of hardware and support personnel for analysis at the U.S. centers would force U.S. physicists to travel to CERN to do physics;
- potentially force the U.S. to renege on international agreements: the U.S. has made commitments to fund equipment to be placed at CERN to ensure prompt and efficient data distribution to the U.S. and other centers (LHC EDGE computing project). This is one of the projects that would need to be significantly descoped with the proposed funding shortfalls.

A second common issue had to do with the operation of grid computing systems. Grid2003, the operation of a full-scale production grid for several months, was viewed as a tremendous success. Both experiments had benefited from Grid2003 operation, running substantial numbers of jobs, and were able to make opportunistic use of resources that were not dedicated to LHC computing. The committee urged both experiments to continue to pursue their strategy of operating heterogeneous grids in some sort of federation, to allow continued exploitation of such resources, particularly at multi experiment and multi science computing centers, and to build on the successes of Grid2003 with plans for the persistent grid infrastructure to be provided by the proposed Open Science Grid.

Finally, it was noted that base program erosion is beginning to have significant effects, especially at universities, who are crucial participants in the Software & Computing projects, and steps should be taken to limit any further erosion.

US ATLAS

In the area of Grids and Facilities, the committee recommends that:

- US-ATLAS should continue to pursue its strategy of interoperability (not mandating LCG-n homogeneity). ATLAS should continue to push interoperability as a high priority within LCG for LCG-n (instead of the completely homogeneous install as in LCG-1)

- US-ATLAS should continue to play an active role in the evolution of Grid 2003 (e.g. Open Science Grid)

- Tier 1 staffing and hardware should be given priority for use of management reserve.

- US-ATLAS should work with BNL management to develop a long range plan for network connectivity.

In the area of Software, the committee recommends that:

- Concerning Funding:

  The agencies should acknowledge that chronic funding shortfalls will lead to a descoping of the project and a loss of the leadership role of the US ATLAS group within ATLAS.

  The US ATLAS software group should clearly outline the consequences of continued funding deficits and do whatever can be done to get a faster ramp up of personnel.

- Providing full support for at least two platforms should be a high priority. This will help QA, debugging and will significantly reduce the risk of being locked in to one platform.

- The existing Quality Assurance (QA) plan should be fully executed. In particular, code reviews should be held regularly.

- The overall ATLAS management and the US ATLAS group should do whatever is possible to resolve the database coordination problem.

- The US ATLAS group should plan for a greater user support load than is currently anticipated, in particular following the coordinated test beam (CTB).

- The US ATLAS management needs to resolve the High Level Trigger (HLT)/Offline management situation, possibly finding a way of managing the projects centrally.

- The US ATLAS group is encouraged to continue the involvement in ARDA (A Roadmap for Distributed Analysis). The group should work very closely with the US CMS group to further a common solution to be adopted by LCG.

In the area of Project Management, the committee made the following recommendations:

- US ATLAS S&C is important to the overall success of ATLAS and to the participation of US physicists in data analysis. The program has made impressive progress since last year. Failure of the program to receive the planned funding would place success in jeopardy. Therefore, we recommend that the funding agencies either meet the profile or formally recognize that the US will be unable to meet all its commitments.

- We recommend that US ATLAS management complete the draft Research Operations Management Plan, and that the plan include details about Software and Computing .

- Efforts to strengthen communication with CERN should continue.

- US ATLAS Software and Computing should continue to be very wary about increases in scope as software developers are heavily committed and important projects have been deferred. The process of changing scope might benefit from becoming more formal.

- The erosion of base program support to the National Labs and the universities has been very damaging. This funding should be increased if at all possible.

US CMS

In the area of Grids and Facilities, the committee makes the following recommendations:

- US-CMS should continue to pursue its strategy of interoperability (not mandating LCG-n homogeneity).
  - CMS should push to make interoperability a high priority within LCG for LCG-2 (instead of the completely homogeneous install as in LCG-1)

- US-CMS should continue to play an active role in the evolution of Grid 2003 (e.g. Open Science Grid)

In the area of Software, the committee recommends that:

- US CMS should resist the temptation to backfill Software Professional positions with Physicists unfamiliar with modern Software Development

- Should consider dangers of not being fully integrated with SEAL

- Providing full support for at least two platforms should be a high priority

- Manpower liberated from the Calorimeter Framework development should remain within the Core Application Software (CAS) project

In the area of Project Management, the committee recommends that:

- Since the US CMS Software & Computing Program has a defined scope that is appropriate for the US as a major participant in CMS, it is important to maintain the funding profile to allow this program to continue at its agreed-to level. Because the LHC and CMS appear to be holding to their schedules, Software and Computing deliverables can no longer slip without serious consequences. Further budget cuts will lead to reneging on US commitments and will jeopardize US participation in physics.

- As the software and computing effort based in the US expands, strong communication with international CMS will continue to be very important. Continued presence of some software personnel at CERN is important. We also recommend that US CMS S&C invite software personnel from CERN to spend time in the US. Worldwide CMS meetings on software and computing could also occasionally be held in the US.

- US CMS S&C management should develop a way to present clearly the effects of milestone delays and reduced functionality on the overall goals of the program. We also recommend that US CMS S&C management maintain a table of milestones showing not only their current projected completion dates but also how those dates have evolved. The current WBS structure that includes descriptions for milestone changes is also valuable and should be continued.

- We recommend that the US CMS S&C continue to improve its contingency planning process. The goal is to be able to react more quickly to possible changes in funding.

The committee feels that it would be useful if the project management plans for both experiments for the current era of research project were written and approved.

- Changing totals, profiles and arrival of funds at the end of the fiscal year greatly complicates the management task. Flexibility from the agencies in shifting support for tasks in different parts of the project is a helpful tool for addressing this.

Table of Contents

1 Introduction

This report is the product of the DOE/NSF review of U.S. LHC Computing Projects held at Fermi National Accelerator Laboratory (Fermilab) on 13-16 January 2004. The review was charged with examining the technical scope, cost and schedule baselines and the project management of these efforts, focusing on the near-term (next two years) plans of both collaborations where plans are more concrete in developing software and user facilities for the LHC experiments, ATLAS and CMS, that are scheduled to begin taking physics data in 2007.

Nine outside experts reviewed presentations made by both collaborations on their individual projects. The evaluations of the experts are contained in this report. At the review, they asked many questions of the presenters and provided many recommendations to the collaborations and to the agencies.

All the outside experts were selected jointly by officers of the Division of High Energy Physics of the US Department of Energy and officers of the EPP Program in the Division of Physics of the National Science Foundation.

In addition to plenary sessions, three breakout sessions were planned in which more detailed presentations of a highly technical nature were made to three sub-committees, each consisting of various members of the review committee. The sub-committees were concerned with: (a) Grids and Facilities, (b) Software, and (c) Project Management.

Observers from the funding agencies were present at all meetings and participated freely both in the open sessions and in the executive sessions. The review was chaired by Moishe Pripstein of the DOE with much assistance from Irwin Gaines, Glen Crawford and Saul Gonzalez of the DOE and Jim Whitmore of the NSF. Ruth Pordes, Michelle Gleason, and Mark Leininger of Fermilab provided invaluable local support.

The charge to the reviewers is shown in Appendix A. The review committee was composed of experts in computing for high-energy physics and related fields, and the committee membership is detailed in Appendix B. The agenda for the review is shown in Appendix C. Separate presentations were made for the U.S. ATLAS and U.S. CMS computing efforts totaling one-and-a-half days for each experiment. Additional material prepared by US ATLAS and US CMS, including cost and schedule tables, status of milestones, and slides of many of the talks, may be found at the experiment web sites for the software and computing reviews. These are:

http://www.usatlas.bnl.gov/computing/meet/0401AgencyReview/

http://www.uscms.org/sandc/reviews/doe-nsf/200401

This report, including its recommendations, represents the views of the committee members on issues raised during the review, but it does not attempt to portray the personal opinions of every reviewer nor to provide a comprehensive summary of all issues related to LHC computing efforts. It is intended as a compendium of expert advice to the funding agencies, and to the U.S. and international collaborators

on the ATLAS and CMS experiments, on how best to achieve the goals of the software and computing projects.

2 Program Overview

CMS and ATLAS will be large, general-purpose detectors for observation of very high energy proton-proton collisions at the Large Hadron Collider (LHC), now under construction at CERN, the European Laboratory for Particle Physics, near Geneva, Switzerland. The LHC will be the highest energy accelerator in the world for many years following its start of operation in 2007. It will provide two colliding proton beams, circulating in opposite directions, at an energy of 7 trillion electron volts (TeV) each, almost an order of magnitude more energy than presently achieved at the Tevatron (1 TeV per beam), at Fermi National Accelerator Laboratory (Fermilab) outside Chicago.

The two large detectors will measure and record the results of the most interesting proton-proton collisions. They will be among the largest and most complex devices for experimental research ever built, and the events that they record are expected to point to exciting, even revolutionary, advances in our understanding of matter and energy. The large increase in energy over that presently available may well lead to an understanding of the origin of mass and the discovery of new types of sub-atomic particles.

The US scientific community strongly and repeatedly endorsed US involvement in the LHC program. Numerous groups of US scientists at universities and national laboratories, historically supported by both the Department of Energy (DOE) and the National Science Foundation (NSF), expressed great interest in the potential physics of the LHC, and in 1994 they tentatively joined the international collaborations designing the CMS and ATLAS detectors. In 1996, DOE and NSF formed the Joint Oversight Group (JOG) to coordinate and manage these efforts and to negotiate an appropriate US role in the LHC program.

In December 1997, the heads of the DOE, NSF and CERN signed an agreement on US participation in the LHC program. This agreement was further detailed by the Experiments and Accelerator Protocols signed later that month, committing the US to spend a total of $531 million on LHC construction projects, with $200 million of that for aspects of the accelerator, and the remainder to support the efforts of the US high energy physics (HEP) community in the construction of the two large detectors. The US efforts on detectors were formalized into construction projects with baselines established in 1998.

US physicists are participating in many aspects of the detectors, including important management roles. With more than 300 physicists from over 30 US universities and four national laboratories working on each of the two large detectors, the US groups comprise over 20% of the international collaborations, and the US groups plan to provide a comparable portion of each detector.

As with past large detector projects, the LHC research program, including the computers and software needed for the physics data analysis, was not made part of the detector construction projects. However, the US LHC research program must be successful if the US HEP community is to reap the scientific benefits of the US investment in the LHC. In addition, the international scientific community is depending on the US to contribute its share of the collaborative effort.

With the construction projects for both of the large general purpose detectors and the accelerator well underway, the Joint Oversight Group, in November 2000, held the first "baseline" review and assessment of the formal organization of the US LHC Research Program, including software and

computing projects that will be central to generating the physics results over the lifetime of the experiments.

Subsequent comprehensive reviews of the US LHC software and computing projects were undertaken in November 2001 and January 2003, concentrating on technical, cost, schedule, and management plans for the near-term efforts (up through the following 2 years) of both collaborations in developing software and user facilities for ATLAS and CMS.

The US LHC Research Program is a joint effort of DOE and NSF, utilizing the oversight structures established for the US LHC Construction Project, as detailed in the DOE/NSF Memorandum of Understanding concerning US participation in the LHC program, as modified to include the US LHC Research Program. In particular, this report is the result of the third formal "baseline" review of the Software and Computing Projects of both US ATLAS and US CMS. This review was conducted in a manner analogous to the DOE/NSF reviews of the US ATLAS and US CMS Detector Construction Projects. However, due to the dynamic nature of the software and computing fields, and the rapidly evolving technology, we do not expect that complete long-term (5-year) baselines could or should be set at this time. The reviewers were therefore asked to evaluate the detailed technical, cost, schedule, and management plans for the near-term (2-year) program efforts, and to evaluate the technical accomplishments in the first full year of baselined operations.

3 Grids and Facilities

US ATLAS - Grids and Facilities

Introduction

The USATLAS project has made considerable progress in the last year in the growth and operation of their facilities and in the maturing use of the grid. The current rate of progress should lead to a successful system being in place in time for detector commissioning.
Overview presentations for this topic were made to the review committee during the first day of plenary talks. Additional grid related material was also presented to some members of the committee during a meeting at Fermilab on a future Open Science Grid, held immediately before the US LHC Computing Review. The break out session on facilities and grids was primarily question and answer with some presentation. The material presented was very good, and the answers to questions were frank and reflected a good understanding of the challenges and plans for solutions.

User Facilities

The USATLAS Tier 1 facility at Brookhaven is now being built out after a long freeze, with a tripling of CPU planned in 2004 from 1% to 3% of the eventual capacity. This will represent nearly 10% of the eventual number of compute nodes, which will allow the facility to gain valuable operations and maintenance experience at an interesting scale and complexity level. The facility is also ramping up their personnel, adding 2 FTE years in FY2004 (3-4 hires). Nonetheless, through 2006 the Tier 1 center will not be significantly larger than the USATLAS Tier 2 centers. It is important that the collaboration begin to invest in technology evaluations, support personnel and hardware. In the past, such investments have been sacrificed in favor of software development. While this made sense in the past, it is time to begin to shift this emphasis. In particular, staffing at the Tier 1 center is currently too tight to address many important tasks at a healthy level, such as disk system evaluations (like dCache) or interacting with external entities such as grid projects.
The Brookhaven facility continues to pursue a plan to have all data on disk rather than on tape. There is strong motivation for this, as it will allow a full sweep through the data in 24 hours as opposed to 1 month if all the data had to be accessed from tape.
The network connection of the center, shared with RHIC computing facilities, is OC-12. This will be saturated by the upcoming Data Challenge 3 (DC3) for ATLAS. This is a concern.
By making opportunistic use of resources, USATLAS contributed more than their fair share to ATLAS Data Challenge 1 (DC1).
Two of the three prototype Tier 2 centers are upgrading by approximately 50%, and currently the capacity of the Tier 2s considerably exceeds that of the Tier 1. Selection of the Tier 2 sites will start this year.
The Tier 1 and 2 centers have sufficient capacity for Data Challenge 2 (DC2).

Grid Evolution

USATLAS has been a major participant in the Grid 2003 demonstration project. Grid 2003 was conceived as a 6 month project leading up to the SuperComputing 2003 conference, with specific goals for scale (number of processors, number of applications) and efficiency (% of jobs successfully

completed). Resources were contributed from a number of sources, including the LHC projects and the U.S. grid projects. The scale of the assembled grid provided important experience in exploiting a (prototype) production grid, and allowed testing of grid production running earlier than would have otherwise been possible. In this it was a resounding success.
Grid 2003 helped to produce a healthy interaction with other grid collaborations, including iVDGL, PPDG, CMS, and others.
Typical Grid 2003 use of the aggregate of ~2700 CPU's was 40%-50% (the balance being used by a large mix of local jobs at each site). There was not enough monitoring in this prototype to unfold the real efficiency of the use of this grid. Most of the use of this grid was in "production" work and not interactive analysis, which will produce far greater demands upon a grid. Although the original demonstration purpose has been fulfilled, Grid 2003 continues to be run productively, serving both as a large scale test platform and a source of production cycles. Discussions are now underway among the participants for follow on work, possibly including a future proposal called the Open Science Grid.
Grid software is maturing (with a long way yet to go). Of particular note, grid monitoring is making good progress, with MonaLisa providing a useful production tool. The emerging use of SRM (Storage Resource Manager) demonstrates a component approach to the grid in which the component interface is specified, and independent implementations (mapped to local storage management systems) are able to interoperate.
So far, USATLAS has used its iVDGL funding for grid tool development (PACMAN, integration and packaging of ATLAS software), and has leveraged personnel from USATLAS institutions for Tier 2 support.
ATLAS needs to continue to espouse an interoperability strategy for the grid, based upon standardized interfaces (like the SRM) instead of homogeneous everything (single operating system, identical software on all nodes). LCG problems (homogeneity requirements) currently preclude the opportunistic use of non-dedicated resources within the LHC grid. Discussions are underway to improve this situation, and the next LCG release is expected to be lighter weight and move in the direction of an interoperable approach. This effort has been too manpower constrained both at CERN, and within ATLAS to provide a robust interaction with the LCG project.
ATLAS has made some significant progress in developing grid software to help them exploit multiple grids (Windmill, Don Quijote, and Capone). The plan is to use this software in DC2, although the schedule is tight.

Potential Impact of Additional Funding Cuts

The committee was asked to consider how well US ATLAS was prepared to deal with potential additional funding cuts, and there is no question that cuts of the scale discussed (15%) would be difficult. The USATLAS project provided well developed plans for how such cuts would be absorbed. Manpower cuts would be painful: it would impact ongoing hardware evaluations, LCG integration work, and other software component evaluations; it would also impact user support and the overall quality of operations.

Recommendations

- USATLAS should continue to pursue its strategy of interoperability (not mandating LCG-n homogeneity).
    - ATLAS should push to make interoperability a high priority within LCG for LCG-2 (instead of the completely homogeneous install as in LCG-1)
- USATLAS should continue to play an active role in the evolution of Grid 2003 (e.g. Open Science Grid)

- Tier 1 staffing and hardware should be given priority for use of management reserve.
- USATLAS should work with BNL management to develop a long range plan for network connectivity.

US CMS - Grids and Facilities

Introduction

The USCMS project has made considerable progress in the last year in the growth and operation of their facilities and in the maturing use of the grid. The current rate of progress should lead to a successful system being in place in time for detector commissioning.
Overview presentations for this topic were made to the review committee during the first day of plenary talks. Additional grid related material was also presented to some members of the committee during a meeting at Fermilab on a future Open Science Grid, held immediately before the US LHC Computing Review. The break out session on facilities and grids was primarily question and answer with some presentation. As in previous years, the material presented was very good, and the answers to questions were frank and reflected a good understanding of the challenges and plans for solutions.

User Facilities

The US CMS user facilities are being upgraded at a reasonable pace. During 2003, the Tier 1 center roughly doubled in size to about 200 boxes as part of the preparation for running Data Challenge DC04. This represents of order 10% of the planned box count at turn-on, and thus represents a reasonably sized system for scaling tests at this point in the project, especially in light of tight budgets. The Tier 1 capability will again double in 2004, servicing both physics users and preparing for DC05.
A modular (cell) architecture has been established for the build-out of the Tier 1 center. This approach aggregates a rack of compute nodes with an ethernet switch and a file server into a replicable unit. Each file server is connected via a fiber channel switch to an array of disks. This approach appears robust and scalable, at a slight cost premium for fiber channel. Software configuration and cluster management is via the ROCKS package.
Progress has been made on securing adequate network bandwidth for LHC operations. In particular, Fermilab has negotiated and committed funds for a connection to StarLight to augment its ESnet connection. This will provide good bandwidth to Europe, at least 1 Gb/s, possibly 10 Gb/s.
Progress has been made on the interoperability of security models between Kerberos and PKI, as demonstrated by the Grid 2003 project described below, and this topic, while still requiring additional development, no longer remains a serious concern.
Staffing remains tight at the Tier 1 center. Staffing was increased somewhat during 2003, but hiring was delayed until late in the fiscal year due to internal Fermilab priorities (re-allocation of existing staff as a priority over new hires). The slow growth in staff was a contributing factor to a slip of 3 months in the delivery of DPE 1.0 (distributed production environment). Staffing should increase an additional 2.5 FTEs in 2004 (with 2 positions open at the time of the review), with an even larger increase in the number of people as hires are staggered into the fiscal year.
The planned 2004 pilot Tier 2 was deferred (funding constraint), but the three prototype centers provided good operating experience as well as increased in capacity. The process of selecting the final Tier 2 centers will begin this year.

Grid Evolution

USCMS has been a major participant in the Grid 2003 demonstration project. Grid 2003 (aka Grid 3) was conceived as a 6 month project leading up to the SuperComputing 2003 conference, with specific

goals for scale (number of processors, number of applications) and efficiency (% of jobs successfully completed). Resources were contributed from a number of sources, including the LHC projects and the U.S. grid projects. The scale of the assembled grid exceeded the milestones of CMS's grid developments, and allowed testing of grid production running earlier than would have otherwise been possible. In this it was a resounding success.

Balancing this appraisal is the fact that it was yet another heroic effort, requiring the efforts of a large number of talented people at the various contributing sites, and only 75% of the jobs ran to completion (which was the goal set in June 2003). Grid 2003 helped to produce a healthy interaction with other grid collaborations, including iVDGL, PPDG, ATLAS, and others.

Typical Grid 2003 use of the aggregate of ~2700 CPU's was 40%-50% (the balance being used by a large mix of local jobs at each site). There was not enough monitoring in this prototype to unfold the real efficiency of the use of this grid. Most of the use of this grid was in "production" work and not interactive analysis, which will produce far greater demands upon a grid. Although the original demonstration purpose has been fulfilled, Grid 2003 continues to be run productively, serving both as a large scale test platform and a source of production cycles. Discussions are now underway among the participants for follow-on work, possibly including a future proposal called the Open Science Grid.

Grid software is maturing (with a long way yet to go). Of particular note, grid monitoring is making good progress, with MonaLisa providing a useful production tool. The emerging use of SRM (Storage Resource Manager) demonstrates a component approach to the grid in which the component interface is specified, and independent implementations (mapped to local storage management systems) are able to interoperate.

In addition to the Grid 2003 work, U.S. CMS has a small number of nodes running the CERN provided LCG 1.0 (LHC Computing Grid) software. Unfortunately LCG 1.0 required a homogeneous (heavy weight) installation, which prevented these nodes from being integrated with the rest of the CMS and Grid 2003 resources.

CMS needs to continue to espouse an interoperability strategy for the grid, based upon standardized interfaces (like the SRM) instead of homogeneous everything (single operating system, identical software on all nodes). LCG problems (homogeneity requirements) currently preclude the opportunistic use of non-dedicated resources within the LHC grid. Discussions are underway to improve this situation, and the next LCG release is expected to be lighter weight and move in the direction of an interoperable approach. This effort has been too manpower constrained both at CERN, and within CMS to provide a robust interaction with the LCG project.

CMS software releases include a test suite to validate an installation at a new site, including a checksum on output files, a commendable approach.

For FY2004, U.S. CMS is planning to study the use of filtering at the Tier 0 center as a way of improving the quality of data sent to the U.S. during operations (Edge Computing). This appears to be a worthwhile strategy with potential gains for the physics in the U.S. This new collaborative work with CERN will require both manpower and the purchase of hardware to be situated at CERN.

Potential Impact of Additional Funding Cuts

The committee was asked to consider how well US CMS was prepared to deal with potential additional funding cuts, and there is no question that cuts of the scale discussed (15%) would be difficult. The U.S.CMS project did not have well developed plans for how such cuts would be absorbed, and so the following assumes that the User Facilities and Grid developments would be cut proportionately (15%). Manpower cuts would be painful: it would impact ongoing hardware evaluations, LCG integration work, and other software component evaluations; it would also impact user support and the overall quality of operations.

One possible area of savings is in the Edge Computing hardware: one could pursue a staged approach, buying ½ of the hardware now to support all of the R&D, and procuring the rest later for the final scaling measurements. This would not seriously impact progress on the Edge Computing task, and would be a better strategy than the complete cut that was proposed by US CMS management as a cost savings measure (far better to proceed at ½ the hardware level than to drop the entire task).
The Tier 1 hardware could be scaled back; this cut would impact DC05 if the shortfall was not made up for early in FY2005. The reduction of resources in FY2004 would impact preparation of the physics TDR, but this could perhaps be recovered by opportunistic use of other resources.

Recommendations

- US CMS should continue to pursue its strategy of interoperability (not mandating LCG-n homogeneity).
    - CMS should push to make interoperability a high priority within LCG for LCG-2 (instead of the completely homogeneous install as in LCG-1)
- US CMS should continue to play an active role in the evolution of Grid 2003 (e.g. Open Science Grid)

4 Software

US ATLAS - Software

Findings

The detector construction is nearly over; this year is the first during which funds for US ATLAS come from the M&O and Software and Computing Project. During this phase of the project, the budgetary allocations for M&O and Computing are decided by the US ATLAS Project Manager, in consultation with the various Level 2 managers. Consequently, there will be a natural tension and competition for constantly diminishing resources for the computing effort.

US ATLAS was directed by the funding agencies to provide an analysis and prioritization of their tasks for FY2004 in the event that the agencies must make a 15% rescission in funds compared to the funding guidance provided at the previous review. US ATLAS did an excellent job of providing clear and lucid presentations and discussions of this scenario for which they are to be commended.

Since the last review, US ATLAS have completely reviewed and revised their Work Breakdown Structure (WBS). They have adopted new schedule analysis tools, such as MS Project, that provide them with better management tools and visibility.

The US ATLAS Level 2 software management has been changed since the last review. This transition appears to have been successfully completed and the new personnel are in place.

The US ATLAS software group contributes to the overall ATLAS core software effort and it provides subsystem software related to the hardware responsibilities of the US ATLAS group. In addition, it participates in the LHC Computing Grid project (LCG). Members of the US ATLAS software group have taken on key responsibilities within ATLAS:

- D. Quarrie (LBNL) is “Software Project Leader”,
- T. Lecompte (ANL) is “Planning and Resources Coordinator”,
- D. Malon (ANL) is “Offline Data Management Coordinator”,
- S. Rajagopalan (BNL) is “Liquid Argon Software Coordinator”,
- S. Goldfarb (U. Michigan) is “Muon Software Coordinator”,
- T. Wenaus (BNL) is “LCG Application Area Coordinator”.

The contributions of the US ATLAS software group are concentrated in these areas:

- Database management,
- Control framework (ATHENA),
- Persistency (POOL and ATHENA/POOL integration).

In FY2003 the US ATLAS software group contributed 6 FTEs out of the total 30 FTEs required for the ATLAS core software project and 3 FTEs out of the total 44 in LCG.

The overall international ATLAS software effort is 37% short of the required/planned-for complement of 164 FTEs working on ATLAS software across the entire collaboration. The US ATLAS component is about right, but is stretched thin due to the overall chronic lack of full manpower within the

collaboration. The international collaboration management is working with other members to correct the situation as rapidly as possible. Nonetheless, the US ATLAS staffing of the group is very lean. This is exemplified by the fact that the database group has never reached the planned level of 6.5 FTEs. Currently it stands at 3.6 FTEs. In spite of this, good progress has been made in key areas:

- The initial integration of the LCG persistency solution, POOL, has been completed. This is a prerequisite for the upcoming data challenge 2 (DC2).
- A new geometry kernel has been developed at the University of Pittsburgh. This provides a consistent geometry description across the entire software suite.

The group would like to maintain its leadership roles in the areas of data management, event store, and detector description. The group also wants to maintain the support for analysis tools. The group also participates in the new ARDA (A Roadmap for Distributed Analysis) project within LCG. However, constant funding pressure forces the group to push back critical tasks into future years with little prospect of funding to recover them. For example, there is currently little work on code hardening and no work on tag databases. The committee was told that the developers are seriously over-worked. The US Database has brought in new members and is bringing them up to speed.

Only one OS platform, namely Redhat Linux 7.3 on INTEL processors, is being supported.

The area of databases is well covered within the offline community of ATLAS due to the strong role of the US ATLAS software group. There is, however, the danger of the proliferation of different and possibly incoherent solutions within the larger ATLAS context due to the absence of overall coordination.

There is also a large functional overlap between the High Level Trigger (HLT) and the Offline software. This concerns the common framework, the event data model and a number of infrastructure components. The primary differences arise from the need of the HLT software to operate in a multi-threaded environment, the seeded mode reconstruction and from the critical performance requirements. Nevertheless, the HLT software is managed separately. The US ATLAS software group would like to change that and manage the projects centrally thus building a coherent US role. Discussions are currently ongoing inside the US ATLAS group to see how this can be achieved.

Only a small staffing increase of 1.5 FTEs is covered by the funding guidance in FY2004. In FY2005 the guidance only covers a flat staffing level since support for some people through GRID projects ends and these people will have to be supported through program funds. Part of this situation is exacerbated by their so-called “University Problem” – US ATLAS is staffed primarily by senior personnel from three US DOE national Laboratories (BNL/ANL/LBNL). While such skills are required to provide the long-term support and focus during the construction of the US ATLAS software infrastructure, the demographics leave them lacking in scientific (postdoctoral, graduate student) personnel at the universities that will be needed and important to fulfilling the promise of scientific exploitation in the US of the ATLAS science. The plan is to ramp up an additional 7 FTEs at the universities. Some progress was reported this year. Further, it is expected that only one additional hire will be made at the national laboratories.

The next major milestones are DC2 and the coordinated test beam (CTB). The goal of DC2 is to process 10 M events over a period of 10 days. This corresponds to 10% of the expected final dataflow for a period of 10 days. In the CTB, a 1/8 wedge of the ATLAS barrel will be placed inside the H8 beam area at CERN. This requires a complete software infrastructure almost identical to the final one except for the geometry. Results from DC2 and CTB are important ingredients for the computing

model document to be produced at the end of 2004. It is estimated that the overall core software effort is about six weeks behind schedule, with DC2 coming soon.

Up to now, the software has not been exercised very much. This will change with the upcoming DC2 in 2Q/3Q 2004. There will be major tests of calibration and test stand hardware, leading to competing (conflicting?) needs between DC2 and the CTB: stability vs. need to continue to test and improve. In addition, the load from users requiring support will likely ramp up significantly, further exacerbating the manpower issue.

During FY2003, ATLAS revamped its high level reconstruction and event data model. The DC2 will exercise and use all new software that became available after DC1. Part of this effort includes a full integration with GEANT4 . In addition, ATLAS has fully adopted/integrated the LCG POOL/SEAL persistency model. They are using the SEAL Foundation for the dictionary and have been plugged into the ATLAS software framework.

The US ATLAS group emphasized how little effort has gone into the development of interactive data analysis environment – reconstruction/analysis tools. The US ATLAS product that has started being developed is DIAL. There is also a BNL contribution – Physics Analysis Tools based on a C++/Python implementation.

Observations

The committee commends the US ATLAS software group on its continuing leadership role in the global ATLAS software effort and within the LHC Computing Grid project (LCG). Members of the group hold key positions inside the ATLAS software organization and within LCG. The group provides strong contributions to the overall software and computing effort. These contributions are at an appropriate level considering the US share in ATLAS and should be maintained in the future.

The group is well organized and properly focused. The momentum working up to the next major milestones has been maintained through the transition to a new Level 2 manager. This is exemplified by the completion of major milestones in 2003. These are the integration with POOL and the implementation of a consistent geometry description across the entire software suite. For the latter task, existing knowledge at the University of Pittsburgh has been leveraged. This is one of the very few examples of strong university participation in the project. Generally, the university participation in the project is weak.

The committee observes that the group appears well prepared for DC2 and CTB; there is significant risk in making up the six-week schedule slip. It should be noted that DC2 has been de-scoped and re-phased. It is planned as a 10% dataflow for 10 days but it exercises the processing and storage infrastructure only at a 1% level. The emphasis is on software integration rather than processing and storage. Concerning processing and storage, the committee notes that the insights gained from DC2 will be minimal and that the data challenges in the coming years will have steep increments that are associated with significant risks.

The committee believes that the planning for user support following DC2 and CTB is inadequate. Without appropriation of further personnel the inevitable support load will strongly affect the software developers' productivity and slow down ongoing projects.

Within ATLAS, the offline and online communities have agreed to make use of a common framework, event data model and reconstruction algorithms. A lack of coherence between the HLT and Offline software groups leads to difficulties in resolving different results between online and offline codes and a duplication of effort and it represents a serious schedule risk. The US is a major contributor in both the development of offline infrastructure (under software program funds) and online software infrastructure (U. Wisconsin under construction funds). The efficiency in the use of personnel and other resources within the US would increase and the US participation would be strengthened if these projects could be managed centrally.

A reasonable software quality assurance (QA) plan has been established and was presented to the committee. This plan comprises a multi-level process starting at the level of individual package developers and ending with validation by the ATLAS physics groups. A procedure for bug tracking is in place. However, the committee observes that the plan is not fully executed due to staffing shortfalls. For example, regular code reviews are planned but are not happening.

In general, it appears that due to excessive funding pressures by the agencies, the US ATLAS group under-represents the true risk to the program. This is exacerbated by the continued erosion of the base program, both at the National labs and at the universities. The tight budget also excludes almost any University participation. This will become more and more painful and will eventually result in too few students who can exploit the Physics potential.

Recommendations

- Concerning funding:
  - The agencies should acknowledge that chronic funding shortfalls will lead to a de-scoping of the project and a loss of the leadership role of the US ATLAS group within ATLAS.
  - The US ATLAS software group should clearly outline the consequences of continued funding deficits and should do whatever can be done to get a faster ramp up of personnel.
- Providing full support for at least two operating system platforms should be a high priority. This will help QA, debugging and will significantly reduce the risk of being locked into one platform.
- The existing QA plan should be fully executed. In particular, code reviews should be held regularly.
- The overall ATLAS management and the US ATLAS group should do whatever is possible to resolve the database coordination problem.
- The US ATLAS group should plan for a greater user support load than is currently anticipated, in particular following CTB.
- The US ATLAS management needs to resolve the HLT/Offline management situation, possibly finding a way of managing the projects centrally.
- The US ATLAS group is encouraged to continue the involvement in ARDA. The group should work very closely with the US CMS group to further a common solution to be adopted by LCG.

US CMS - Software

Findings

The LHC and the CMS detector appear to be on schedule for a 2007 start of data taking.

There is clearly a very strong US Contribution to the CMS Core Application Software (CAS) effort. The US team is involved in almost all aspects of the CMS Software development and they have had a profound impact in the CMS development model. This is providing the US with a very strong leadership role in the CMS Software effort. The US CAS Team is stretched very thinly. In the CMS Computing & Core Software project the following roles are undertaken by US participants;

- D. Stickland (Princeton U.): Project Manager,
- Lucas Taylor (Northeastern U.): Deputy Project Manager,
- Lucas Taylor (Northeastern U.): (Acting) Technical Coordinator,
- Lothar Bauerdick (Fermilab): Regional Center Coordination,
- Tony Wildish (Princeton U.): Production Processing and Data Management,
- Stephen Wynhoff (Princeton U.): Reconstruction Manager.

Four other posts in the organization are held by CERN staff and a final one by an Italian. The US also has involvement in the HCAL, muon and heavy ion areas of the CMS Physics & Reconstruction Software project (together with the Reconstruction Manager above which is a cross project position).

LHCC agreed with CMS that they are short of 8 FTEs. The shortfall in manpower is 27% of the total need. However, the US is providing its share as 8 FTEs which is 29% of the total needed. The overall shortage at the international level is leading to reduced effort across the board rather than descoping. This introduces risk. One example that was offered is a scale back of the Iguana maintenance and support effort. The US has been pressing its international partners to provide additional FTEs, but the effort has not borne fruit yet. The scenario that was depicted is one in which everyone is extremely overworked – has been for 3 years and they still have 3 to go.

The software group was not prepared to answer the funding agency request to address a potential FY2004 15% shortfall of NSF funds.

The CMS S&C program is beginning to attract physicists from the Base Program in Universities and Laboratories. This includes physicists who are completing their detector construction and/or development activities and physicists from other experiments. Nevertheless, the overall Base Program continues to shrink.

The US CMS CAS passed the Calorimeter Framework development task responsibility back to the central CMS Organization (which has been assigned to the Physics & Reconstruction area and will probably be performed by US physicists). This potentially frees 1 FTE of effort for other tasks within the Core Application Software area.

The CMS Collaboration has completed the migration from Objectivity/DB persistence to use the LHC Computing Grid's POOL storage framework. This was a large effort provided by US CMS.

The adoption of the SEAL Foundation by CMS has been limited to areas required for the use of POOL.

The CMS collaboration supports the Linux and Solaris platforms. Due to the lack of widespread interest in Solaris they are considering the support of MacOS X (really BSD Unix) instead.

The User Support load on the computing organization has been ramping up with the commissioning of the detector. It is foreseen that this effort will require about 0.25 FTE of effort from all of the CAS developers after FY2004.

For software QA, the CMS collaboration uses various tools. These include two offshoots from the Iguana project: Ignominy & igprof.

US CMS is successfully leveraging synergistic relationships with Grid collaborations in the US.

Observations

The US contribution to CMS software is very fundamental. Any reduction in the US CAS effort would have a large negative effect on CMS software. In addition, the brunt of that impact would disproportionately impact the U.S. physicist's ability to carry out physics from their home institutions, as the areas of software effort that could be most easily de-scoped without sacrificing the whole program are the development efforts that would result in a seamless world wide analysis environment.

The Base Program erosion and the accretion of new Physicists are causing a skill mix change from Software Professionals to Experimentalists or Scientific Programmers. This could cause the loss of key developers and expertise.

Some delays in milestones and de-scoping of functionality have occurred due to manpower shortages.

The SEAL Framework has not been fully integrated into CMS Framework. While this seems to have been the prudent decision at the present time due to manpower shortage, divergence in the future could introduce other problems. At some point, if manpower becomes available, the team would like to go back and replace the Iguana class libraries in the core CMS software with the SEAL equivalents.

The support and validation of more than one platform provides improved code quality and robustness. This would also mitigate any risks from depending on a single vendor. The validation of pre-built executables on a new platform (for example, a newer version of a supported platform) can aid in exploiting resources (either on the grid or shared facilities).

It is not clear if the approximately 2.5 FTEs are adequate for the further expected ramp up of commissioning support needed in future. In addition, it is not clear what the interplay of support needed for/from the User Facility, Physics Analysis Center and Core Application Software is.

There was not much discussion of validation during the review. This does not seem to be an issue currently.

Recommendations

- US CMS should resist the temptation to backfill Software Professional positions with Physicists unfamiliar with modern Software Developer techniques.
- US CMS should consider the dangers of not being fully integrated with SEAL.
- Providing full support for at least two platforms should be a high priority.
- Manpower liberated from the Calorimeter Framework development should remain within the CAS effort

5 Project Management

US ATLAS – Findings and Evaluation

Organization

The organization of the US ATLAS Software and Computing Program has undergone some changes since last year. William Willis remains the Project Manager for US ATLAS. John Huth (Harvard U.) is the Associate Project Manager for Physics and Computing with responsibilities for grids and the interface with external projects. James Shank (Boston U.) is the Executive Project Manager for Physics and Computing with responsibilities for Software and Facilities. Last year, Huth was the sole Associate Project Manager and Shank was his deputy. Ian Hinchliffe (LBNL) is the head of Physics (WBS 2.1). Srini Rajogopalan (BNL) is the head of Software (WBS 2.2). Bruce Gibbard (BNL) and Rich Baker (U. Chicago) are the heads of faciltities (WBS 2.3). The Level 3 managers are also in place. This new organization appears to be working well.

The members of US ATLAS are well represented in the new ATLAS computing organization. The heads of the Software Project (D. Quarrie, LBNL), Data Management (D. Mallon, ANL), and Planning and Resources (T. LeCompte, ANL) are all from US ATLAS, as are many of the sub-project leaders. US ATLAS is providing its share of ATLAS Computing leadership.

Scope

The Work Breakdown Structure (WBS) for US ATLAS Software and Computing has two main elements: Software and Facilities.

Software
The US ATLAS software effort (WBS 2.2) has four main level 3 tasks.

- Core Services (WBS 2.2.2).
- Data Management (WBS 2.2.3).
- Application Software (WBS 2.2.4).
- Software Support (WBS 2.2.5)

The other level 3 task is Coordination.

The software effort is growing, although at a rate limited by funding. Core Services had 3.75 FTEs in FY2003 with a request to increase to 6.5 in FY2004. Data Management had 3.6 FTEs in FY2003 with a request to increase to 6.5 in FY2004. Software support had 1.25 FTEs in FY2003 which is expected to be constant in FY2004.

Facilities

The Facilities effort (WBS 2.3) is responsible for the design, installation and commissioning of the Tier 1 and regional Tier 2 computing centers. It also develops computing systems for production and

analysis. BNL is the Tier 1 site. Currently there are two prototype Tier 2 centers at Indiana University and Boston University. Plans call for a total of five Tier 2 sites eventually. The final Tier 2 site selection is scheduled for this year. There are currently 9 active Tier 3 (institutional) centers. The current effort is ~ 19 FTEs, including only 5.4 paid by the program. The other FTEs come from a variety of sources: grid projects (mainly iVDGL and PPDG) and base-program funded physicists.

The Tier 1 Center at BNL is co-located and operated with the RHIC Computing facility. This has helped reduce FTE needs and there has been substantial benefit from cross use of idle resources. However, if the efforts diverge (e.g. Linux vendor/version), these benefits may decrease.

The Production and Grid Tools WBS task were moved from under Software to under Facilities this year.

Schedule

US ATLAS Software and Computing managers have developed a detailed, scrubbed MS project and WBS description of their S&C tasks with milestones. They have prioritized their tasks and know what is at risk if their funding is cut. The committee commends US ATLAS for developing these project management tools and for setting their priorities and describing them clearly.

The ATLAS Software and Computing schedule is built around a series of data challenges and dates for major deliverables in the WBS and MS project. The next Data Challenge (DC-2) is scheduled for April-July, 2004. Its major focus is validating the computing model.

US ATLAS Software and Computing deliverables are about six weeks behind the original schedule. The WBS planning process has enabled ATLAS to accommodate some of these delays by phasing DC-2 and parts of the work needed for the 2004 testbeam run.

Keeping to the schedule is more important than ever in light of the progress on the LHC machine. The committee was told that the LHC schedule is no longer slipping. Magnet deliveries are keeping up with the required baseline. Counting on further slippage in the LHC schedule would not seem to be a prudent choice.

Cost and Funding

The Executive Project Manager for Physics and Computing presented a profile of Computing and Physics program. The given total from FY2003-FY2008 is $60M. The FY2004 total is $4633K and FY05 $7794K. The decline in base program support has reduced the personnel available to the program from the National Laboratories and the universities. This is a continuing serious problem for ATLAS S&C. Deliverables have slipped and important areas of the project are short of people. For example, US effort in the combined testbeam is minimal and only a few people are looking at the DC2 data. This will leave US physicists without sufficient experience to be major players at the turn-on of the LHC when many of the discoveries will be made. US ATLAS program budget reductions have led to a lack of these funds for many universities, including those with the expertise to provide ATLAS with substantial help.

A possible 15% budget reduction in FY04 would have a serious impact on important software and facilities projects. Grid and framework integration would essentially be stopped and many universities would not be able to participate in Data Challenge 2 (DC2). This would result in the Grid being much less useful for ATLAS, and the universities would not be able to use the data challenges to be ready for commissioning the detector and the initial round of ATLAS physics. It is, of course, in the initial round where most of the discoveries are likely to be made. The committee notes that US ATLAS has a cost profile that matches the funding guidance and has a roughly 25% management reserve up to FY 2007. The committee was told that with the current plan, US ATLAS Software and Computing will need more in FY 2008 than they would get if the 25% management reserve is preserved.

Project Control

Scope is determined by consultation with the Level 2 and 3 managers and with international ATLAS software coordinators. While the committee believes that scope is better controlled now, the process still seems informal and too few software agreements are in place. Scope control is still important as the committee was told that software developers are overworked and at risk for burnout.

It appears to the committee that project control for US ATLAS software and computing is not well defined at this time. The draft Research Operations Management Plan contains very little about Software and Computing.

Recommendations

- US ATLAS S&C is important to the overall success of ATLAS and to the participation of US physicists in data analysis. The program has made impressive progress since last year. Failure of the program to receive the planned funding would place success in jeopardy. Therefore, we recommend that the funding agencies either meet the profile or formally recognize that US ATLAS will be unable to meet all its commitments.

- We recommend that US ATLAS management complete the draft Research Operations Management Plan, and that the plan include details about Software and Computing .

- Efforts to strengthen communication with CERN should, of course, continue.

- US ATLAS Software and Computing should continue to be very wary about increases in scope as software developers are heavily committed and important projects have been deferred. The process of changing scope might benefit from becoming more formal.

- The erosion of base program support to the National Labs and the universities has been very damaging to US ATLAS. This funding should be increased if at all possible.

US CMS – Findings and Evaluation

Organization

The organization of US CMS Software and Computing (USCMSSC) has been stable over the last year. The transition to operation within the US CMS Research Program is happening smoothly. L. Bauerdick continues as Level 1 Manager. At the time of the last review, he was also the acting Level 2 manager for User Facilities, but this position has now been filled by I. Fisk (Fermilab). Fisk was formerly the Level 2 Manager for Core Application Software – this position is now held by R. Clare (University of California, Riverside). There had been plans to add a third Level 2 manager for Grid activities, but US CMS Software and Computing has elected to stay with the current structure of only two Level 2 managers.

Communication with the international CMS Computing organization is facilitated by continued strong US presence in the Computing and Core Software (CCS) group. D. Stickland (Princeton University) continues as CCS Project Manager. Other US CMS members of the CCS Steering Committee include L. Taylor (Northeastern University), L. Bauerdick (Fermilab), T. Wildish (Princeton University) and S. Wynhoff (Princeton University). US physicists are represented on many CMS Software and Computing Boards and hold management positions at various levels in many projects.

The US CMS Software and Computing managers have a new draft management plan that describes USCMSSC's operation within the US CMS Research Program. This plan will be combined with a plan for Maintenance & Operations (M&O) and some unifying text to form a complete management plan for the US CMS Research Program.

Scope

The Work Breakdown Structure (WBS) for USCMSSC has two elements: Core Application Software and User Facilities.

Core Application Software
The USCMSSC core software effort is focused on three main areas:

- Architecture frameworks Toolkits and Reconstruction Software
- Visualization and Analysis.
- Distributed Computing.

USCMSSC is also committed to supporting physicists in both the development of reconstruction code and data analysis. There are 2.2 FTEs assigned to user support.

Core Application Software is managed as a Level of Effort contribution to the CMS Core Software. Currently, the US contribution to the CCS effort (software professionals only) is ~ 9 FTEs spread over 14 people. This amounts to about 29% of the required CCS effort. Overall, the CCS effort has a deficit of 27%, but the CCS manager indicated that the US is doing its share. For comparison, the same chart shown at the January, 2003 review did not show any deficit but showed 29% of the CCS effort coming from the LHC Computing Grid (LCG).

User Facilities

The User Facilities effort is responsible for the design, installation and commissioning of the Tier 1 and regional Tier 2 computing centers. It also develops computing systems for production and analysis. Fermilab is the Tier 1 site. Currently there are three prototype Tier 2 centers at Caltech, UC San Diego and University of Florida. Plans call for a total of five Tier 2 sites eventually. The final Tier 2 site selection will probably be done this year. The current effort is ~ 17.25 FTEs, including ~3 FTEs from the PPDG and iVDGL projects.

There are plans for a Physics Analysis Center and Virtual Control Room at Fermilab. This is funded separately from the US CMS Research Program M&O rather than from Software and Computing.

Cost and Funding

The USCMSSC Management Plan (draft v2.06) states that the cost objective is $70M, integrated from FY 2002 to FY 2008. At the January, 2003 review, the Level 1 manager presented a "bare bones" cost profile integrating to ~$67M for FY 2003 – FY 2008. The cost profile presented at this review (January, 2004) was unchanged from the profile presented a year ago.

The Level 1 Manager presented a profile of Research Program guidance. The sum is $50.8M (NSF) + $93.6M (DOE) for a total of $144M from FY 2002 – 2008. In the current plan this is split $74M for S&C and $70M for M&O. There is a procedure in place to determine S&C's share of the Research Program funding each year. The procedure involves consultation with management personnel and committees. The complexity of the program makes it difficult to react quickly to possible funding shortfalls.

A large amount of the S&C funding goes to salaries. This limits flexibility in responding to possible funding shortfalls. Flexibility in use of funding has been important to help the S&C program when funding arrives late in the year. This flexibility is expected to decrease in the future.

USCMSSC management was reluctant to respond to the possibility of additional 15% cuts in funding in FY 2004 as they felt that their plan was already "bare bones". When pressed, they suggested the following possibilities:

- Cut the M&O program by 15%. This will lead to an estimated delay of three months.
- Cut the US CMS participation in the Open Science Grid.
- Cut the US Edge Computing contributions to CERN. This would compromise US access to LHC data and break or delay the agreement reached with the Computing Resource Review Board.

Schedule

The USCMSSC schedule is built around a series of data challenges and submission dates for Technical Design Reports (TDR). The next Data Challenge (DC04, 5% Data Challenge) is scheduled for March - April, 2004. The Computing TDR is scheduled to be submitted in December, 2004, with the Physics TDR following about a year later.

A year ago, the review committee noted, "The committee cannot comment on how well the software and computing milestones are being met, as we were not shown a comparison of prior year milestones and actual completion dates." This was also the case for the 2004 review. Although the committee was told that milestones had been achieved in the past year (and sometimes when they had been achieved), there was no reference to any plan provided. This makes it difficult to assess progress relative to the schedule plan.

Nevertheless, it is clear that there has been significant technical progress over the past year and the group is well on its way to achieving the important DC04 milestone with only a minor delay despite a number of setbacks in external deliverables. The essential milestones for FY 2003 have been achieved and an evaluation of lessons learned has been done. Keeping to the schedule is more important than ever in light of the progress on the LHC machine and the CMS detector. The committee was told that the LHC schedule is no longer slipping. Magnet deliveries are keeping up with the required baseline and US CMS experiment components are on schedule.

Project Control

The USCMSSC Management Plan describes various groups in place for management oversight. Fermilab has a Project Management Group that meets regularly. There is also a Software and Computing Oversight Panel (SCOP) that meets once or twice a year. NSF/DOE typically schedules an annual review of the project with a shorter status report meeting scheduled mid-year. Written reports are submitted quarterly.

The Management Plan includes a description of a change control process. It also includes appendices that describe how the baseline was developed. Software & Computing work at US CMS universities is defined through Memoranda of Understanding (MOUs) and Statements of Work (SOWs).

Recommendations

- Since the US CMS Software & Computing Program has a defined scope that is appropriate for the US as a major participant in CMS, it is important to maintain the funding profile to allow this program to continue at its agreed-to level. Because the LHC and CMS appear to be holding to their schedules, Software and Computing deliverables can no longer slip without serious consequences. Further budget cuts will lead to reneging on US commitments and will jeopardize US participation in physics.
- As the software and computing effort based in the US expands, strong communication with international CMS will continue to be very important. Continued presence of some software personnel at CERN is important. We also recommend that US CMS S&C invite software personnel from CERN to spend time in the US. Worldwide CMS meetings about software and computing could also occasionally be held in the US.
- US CMS S&C management should develop a way to present clearly the effects of milestone delays and reduced functionality on the overall goals of the program. We also recommend that US CMS S&C management maintain a table of milestones showing not only their current projected completion dates but also how those dates have evolved. The current WBS structure that includes descriptions for milestone changes is also valuable and should be continued.

- We recommend that US CMS S&C continue to improve its contingency planning process. The goal is to be able to react more quickly to possible changes in funding.

6 Appendices

Appendix A - Charge to Committee

Charge for Jan 2004 US LHC Computing Review

An independent peer review of the US LHC Software and Computing (S&C) projects will be conducted at FNAL on Jan 13-16, 2004. This review will continue systematic oversight of the US LHC research program.

The scope of this review is to include both the individual US ATLAS and CMS S&C Projects and the common projects which provide software resources to both experiments.
The goal of this review is to assess the current understanding of the scope, cost and schedule for the US LHC S&C projects and the operation of their management structures.
Both US ATLAS and US CMS should present self-consistent project plans targeted to the funding guidance received from DOE and NSF, and separately address how incremental funds could be used (as well as discussing contingency planning that could deal with possibly reduced funding). It is clear, due to the dynamic nature of the software and computing fields, that plans for the next 2 years will be much more concrete than longer term plans, but enough information should be presented to allow the reviewers to judge the adequacy of proposed long term resources.

This year's review should concentrate on understanding the scope of the US projects and the set of milestones that are to be used to assess project progress. In particular, the reviewers should examine how the US Software and Computing projects take on tasks on behalf of the international projects, how they determine their resource needs, and how their scope is matched to funding guidance. They should look at the relationship between the Software & Computing Projects and the overall LHC Research Program, and should understand how resources are allocated and how contingency is managed.

Furthermore, since the projects have been operating for roughly two years since initial baselining, it is appropriate to review achievements of the past year, in comparison with the previous set of stated objectives, and to monitor the level of technical progress.

The charge for this review is to assess:
- The overall scope of the US LHC S&C efforts and their connections to both the international LHC S&C efforts and the CERN LCG project. Is the scope well defined, and are there mechanisms in place to control “scope creep”;
- The risk to US LHC S&C schedule or scope given current funding profiles and overall LHC project schedule, and additional risks from possible reduced funding or increased M&O resource demands;
- The function, scope and structure of the national ("Tier 1") US LHC computing facilities, their relationship to smaller regional and university facilities, and how existing computing centers are being leveraged to provide resources for LHC computing. In particular, are there sufficient resources for the

US Tier I centers to acquire enough hardware resources to play a major role in upcoming data challenges;
- The contributions of the US S&C projects in providing and supporting core and detector specific software components to the international efforts, and the level of leadership from US collaborators;
- The level of integration of computing infrastructure efforts (such as networking and grid computing) into the planning and execution of US LHC S&C projects, particularly the level of control project managers are able to exert. In particular, how are contributions of software from Grid R&D projects managed, and how well are US efforts integrated into overall LCG (LHC Computing Grid) Project efforts;
- The plans of the US collaborations to provide computing resources to their users and their success in integrating US physicists into the software development process;
- US contributions to recent and forthcoming data challenges, both in providing computing facilities and in developing software and managing these challenges;
- Existing and potential common projects which could benefit both ATLAS and CMS; and
- Project Management Plans, organizational structures, adequacy of personnel, contingency planning, and flexibility of each of the US LHC S&C projects.

Appendix B - Members of Review Committee

Moishe Pripstein, Chair		DOE	moishe.pripstein@science.doe.gov
Jim Whitmore	NSF		jwhitmor@nsf.gov
Deb Agarwal		LBNL	DAAgarwal@lbl.gov
Richard Dubois		SLAC	richard@slac.stanford.edu
Steve Gowdy		SLAC	gowdy@slac.stanford.edu
Tobias Haas		DESY	Tobias.Haas@desy.de
Nick Hadley		U Maryland	njhadley@physics.umd.edu
Albert Lazzarini		Caltech	lazz@ligo.caltech.edu
Patty McBride	FNAL		mcbride@fnal.gov
Cathy Newman-Holmes		FNAL	cath@fnal.gov
Paul Sheldon		Vanderbilt	Paul.Sheldon@Vanderbilt.edu
Chip Watson		TJNAF	chip.watson@jlab.org

Appendix C - Review Agenda

CMS

Day 1, Plenary Session, Tuesday January 13, 2004, Location: 1West
Video Streaming available in 1West.

Start of Meeting 8:30am CST

Time	Title	Speaker	Duration
8:00 AM	Video Setup		
8:30 AM	Executive Session (closed, in WH1E)	Panel	30'
9:00 AM	U.S. CMS Research Program [slides]	Dan Green	25'+5'
9:30 AM	U.S. CMS S&C overview [slides]	LATBauerdick	60'+10'
10:40 AM	Coffee		20'
11:00 AM	CMS Software Overview [slides]	David Stickland	35'+5'
11:40 AM	CAS Overview [slides]	Bob Clare	40'+10'
12:30 PM	Lunch		60'
1:30 PM	UF Overview [slides]	Ian Fisk	40'+10'
2:20 PM	Grid2003 [slides]	Ruth Pordes	25'+5'
2:50 PM	Coffee		10'
3:00 PM	Parallel Sessions		150'
5:30 PM	Executive Session (closed, in WH1E)	Panel	30'
6:00 PM	Adjourn		

Day 2, Wednesday, January 14. Location: 1 West
Start of Meeting 8:30 am CST

Time	Session	Duration
8:45 AM	Video Setup	
9:00 AM	CMS answers to questions	120'
11:00 AM	Discussion and committee exec session (WH1E)	90'
12:30 PM	Lunch	60'
1:30 PM	Committee exec session (writing, in WH1E)	120'
3:30 PM	CMS closeout	90'

Parallel Sessions

Core Application Software and User Analysis
Location: 7X (VRVS is in the Forest virtual room)

Fabrics, Testbeds and Distributed Production Grids
Location: Curia II

Project Management, Relationship to External Entities
Location: 1 North

ATLAS

Thursday 15 January 2004

Executive Session (08:30->09:00)

Chair: I. Gaines
Location: FNAL

Plenary (09:00->12:00)

Chair: I. Gaines
Location: 1 West

09:00	**Project Status** (45') (transparencies)	**J. Shank**
09:45	**ATLAS Computing Overview** (30') (transparencies)	**Dario Barberis**
10:15	**ATLAS Software Project overview** (30') (transparencies)	**D. Quarrie**
10:45	**Software Overview** (45') (documents transparencies)	**S. Rajagopalan**
11:30	Coffee	
11:45	**US ATLAS and other Computing Projects** (45') (transparencies)	**J. Huth**
12:30	**Facilities Overview** (45') (transparencies)	**B. Gibbard**
13:15	Lunch	
14:00	**Grid Tools and Services/DC Production** (45') (transparencies)	**R. Baker**

Breakout 1: Project Management (15:00->17:00)

Location: 7th Floor Crossover

Breakout 2: Facilities/Grids (15:00->17:00)

Location: 1-North
Material: more information

15:00	**Grid Production - additional material** (15') (more information transparencies)	**De, Kaushik**

Breakout 3: Software (15:00->17:00)

Location: 1-East

Executive Session (17:00->18:00)

Chair: I. Gaines
Location: FNAL

Plenary Session (18:00->19:00)	**Description:** *Committee questions to ATLAS. Homework questions. All invited* **Chair:** I. Gaines **Location:** FNAL

Friday 16 January 2004

ATLAS questions (08:30->10:00)	**Description:** *ATLAS members answer homework questions.* **Chair:** I. Gaines **Location:** FNAL
Executive Session (10:00->11:00)	**Chair:** I. Gaines **Location:** FNAL
Executive Session (11:00->12:00)	**Chair:** I. Gaines **Location:** FNAL
Closeout (13:30->14:00)	**Chair:** I. Gaines **Location:** FNAL